

ALERTA DE AUDIO FALSO GENERADO POR INTELIGENCIA ARTIFICIAL



<https://x.com/elvoceroecu/status/1904306470528622658?s=48&t=OVpMOMCwyZebSfuaKIOEPO>

ESTE AUDIO ES FALSO Y HA SIDO GENERADO CON INTELIGENCIA ARTIFICIAL. PARA ANALIZARLO INSERTAMOS EL VIDEO EN UN PROGRAMA DE ANALISIS DE AUDIO Y VIDEO DE AI www.resemble.ai PARA ANALIZAR DETALLADAMENTE EL USO DE INTELIGENCIA ARTIFICIAL EN LOS PATRONES DE LA “SUPUESTA VOZ DE JAN TOPIC” Y MANUALMENTE CON EL SCRIPT “DESCRIPT” DE CHATGPT 4.0 LO COMPARAMOS CON EL AUDIO DE UN VIDEO REAL DE JAN TOPIC. COMPARTIMOS EL ANALISIS RESPECTIVO.

INFORME TÉCNICO DE ANÁLISIS DE VOZ

Fecha: 25 de marzo de 2025

Análisis realizado por: Mónica Santos, con soporte técnico de inteligencia artificial

Archivos analizados:

- **Audio 1:** ssttwitter.com_1742952092824.mp4 (sospechoso)
- **Audio 2:** AUDIO-2025-03-25-13-28-19.m4a (referencia original)

Análisis:

- **a. Timbre y tono:**
La voz del primer audio imita el timbre general de la original, pero presenta una textura **plana**, sin las microvariaciones naturales de una voz humana. Esto sugiere una síntesis completa o clonación de voz mediante IA.
- **b. Entonación y emoción:**
El tono emocional del audio sospechoso es **neutro y lineal**, sin inflexiones

propias del lenguaje espontáneo. La voz original presenta variaciones emocionales auténticas, ausentes en el primer archivo.

- **c. Ritmo y pausas:**

El ritmo en el audio 1 es **anormalmente perfecto**, con pausas uniformes y sin errores naturales del habla (muletillas, cortes, repeticiones). Esto indica una generación programada, no espontánea.

- **d. Ausencia de respiraciones:**

En el audio sospechoso **no se detectan respiraciones naturales**, ni sonidos pulmonares ni pausas fisiológicas, lo cual es típico en voces sintéticas.

- **e. Articulación y dicción:**

El nivel de claridad en la pronunciación es excesivo, con **articulación limpia y artificial**. La voz original, en cambio, tiene un ritmo más humano y realista, incluyendo pequeñas imperfecciones.

Basado en el análisis comparativo con la voz original, el resultado estima que el **100% del contenido del audio en el video fue generado o transformado por inteligencia artificial**. No se trata de una simple edición o mezcla, sino de una **síntesis completa de voz**, es decir, el mensaje fue recreado imitando la voz original.

Este tipo de generación se puede lograr con herramientas como **ElevenLabs, Play.ht, Voicery** o **Descript Overdub**, que permiten clonar una voz humana con muestras de audio reales.

Se concluye que el audio `ssstwitter.com_1742952092824.mp4` fue **generado completamente por inteligencia artificial**, utilizando herramientas de clonación de voz.

No se trata de una grabación auténtica ni una edición parcial, sino de una **reconstrucción total del mensaje usando un modelo sintético que imita la voz original**.

Artificialidad en audio/video:

1. Entonación monótona (plana):

- La entonación no refleja emoción.
- No hay cambios sutiles en el tono según el contenido del mensaje (algo que hacemos de forma inconsciente al hablar).

2. Cadencia y ritmo excesivamente perfectos:

- Las frases tienen una fluidez **demasiado constante**, como si estuvieran "programadas".
- Las pausas son **matemáticamente precisas**, no hay vacilaciones ni interrupciones típicas del habla humana.

3. Falta de respiraciones naturales:

- En una grabación humana es común notar pequeñas respiraciones entre frases. Aquí, **la voz es continua**, sin ruidos pulmonares ni pausas de inhalación.

4. Articulación artificial:

- Las palabras están **demasiado bien pronunciadas**, sin los matices o pequeños errores que puede tener un hablante nativo.
- Las consonantes suenan levemente más duras o "limpias", algo típico de voces sintéticas.

5. Ausencia de textura vocal:

- El timbre es correcto, pero **carece de las microvariaciones** que tiene una voz humana por la vibración natural de las cuerdas vocales.
- No hay "aire" en la voz ni texturas orgánicas (como temblores, asperezas, tensión, etc.).

6. Emoción simulada o ausente:

- Si bien la voz puede sonar segura, **no transmite verdadera intención emocional**.
- Esto es especialmente evidente al compararlo con la voz real, donde se nota tono emocional, intención comunicativa y expresión.

Advertencia sobre el uso de IA para desinformación:

- Este caso representa un **claro intento de manipulación digital** mediante herramientas de inteligencia artificial que permiten imitar voces reales con alto grado de fidelidad, pero sin el consentimiento de la persona afectada.
- Alertamos sobre el riesgo creciente del uso malicioso de estas tecnologías para generar **contenidos falsos, manipulados o descontextualizados** con fines de engaño, difamación o propaganda.